A Model of Re-use of E-Learning Content

Paul Libbrecht paul@(activemath.org|dfki.de)

Competence Center for E-Learning, DFKI GmbH and University of Saarland, Saarbrücken, Germany

Abstract. Re-using e-learning content is a solution often proposed against the high-cost of the creation high-quality e-learning material, a solution with a neglectable price. In this paper we propose a content distribution model that addresses the long term evolution of content pieces whose quality raises along with re-uses. A realization of this model is presented for the ActiveMath learning environment platform.

Introduction

E-Learning content, on the one hand, is known to be expensive to realize; on the other hand, its potential for reproduction is much greater than paper-published content. Thus, the idea of long-term development and re-use of content has emerged and has been attempted and studied for example in [1,2,3]. We have found, however, very few studies that address the management of the long-term evolution of content linked to re-use actions such as aggregation, transmission, and publication. That is, re-use seems to be considered as a single shot action whereas a significant quality evolution can be obtained as a result of several re-uses.

This paper advocates the notion of a **content-collection** corresponding to the organization of content projects in **shared directories**, along with the mechanism of **item-inclusion**, and the approach of **semantic content**. Together, these notions allow the project-based maintenance, connecting to author communities, and the assembly and preview actions for the purposes of specific learning experience where re-used content appear as a coherent entity.

This paper starts with an ideal re-use scenario and surveys current practice of re-use. It follows with the various relations of an author to them in the re-use paradigm. Related work and open questions conclude it.

1 A User-Story of Re-use

Let us imagine an author who is assembling the content for next year's course. For this purpose, a new content-collection is set-up. It starts as an inclusion of last year's project but the author wishes to replace parts and does not wish to affect last year's content. So his collection *extends* last year's collection, by referencing it.

The author also wishes to include the interactive exercises of this fellow he exchanged with since last workshop as well as the high-quality real-world examples of a big industrial project he encountered in his professional associations' online community. To be able to evaluate the content elements, he needs to see web-pages that describe the content projects, that point to public demoes and that provides the IPR statements; these information pages allow him to track the online spaces where these content projects happen and document how to include them.

Using the simple *import* facility of his authoring learning platform, he can see the *books* of these content projects and can browse the content within the realm of his own server. When he starts to assemble the contents of the first lessons, he quickly realizes that his students will need more technical instructions for the interactive exercises of the fellow, and also that some of the real-world examples refer to concepts that are formulated differently than the formulations he wishes.

These adjustments are possible for him since he knows how to change the sources of the documents that can be copied and processed by his authoring tools. Doing this, he creates an *appropriated* version of each of the three collections the course he prepares is based on. The content project of this year is first built with just a few *books* that imitate the books of last year and slowly get adapted to incorporate the two new content projects and further explanations.

After the modifications, the author can upload his course to the server of the school. This means uploading all four collections to this server (three base collections and the collection of the new course). Students will see it as a coherent single course. Advanced users of the school's server, e.g. remote teachers, will still be able to click through until the original collections, following an information page with *copyright information*.

Figure 1 presents the view of the working-space of the author, in the central rectangle, with imported collections linked to their external *repositories*, and with a link to the target learning environment.

Because his way of working has clearly identified the derivative nature and the origin of each content collections he is based on, our author can incorporate in the new course, a few months later, corrections to last year's course or revisions to the industrial examples. Similarly, he is able to transmit the enriched technical instructions he wrote so that his fellow considers them for inclusion in his project's repository...

2 Elements of Re-use

By re-use we mean the action of taking an *existing* piece of content with the activity of the creation of content different than the original one.

The authoring activity naturally happens within the context of several **spaces**, such as the composition space (world of the editor) and the preview space (a small learning platform) and it may lead to a publication space (for other authors) or a staging space (for the learners).

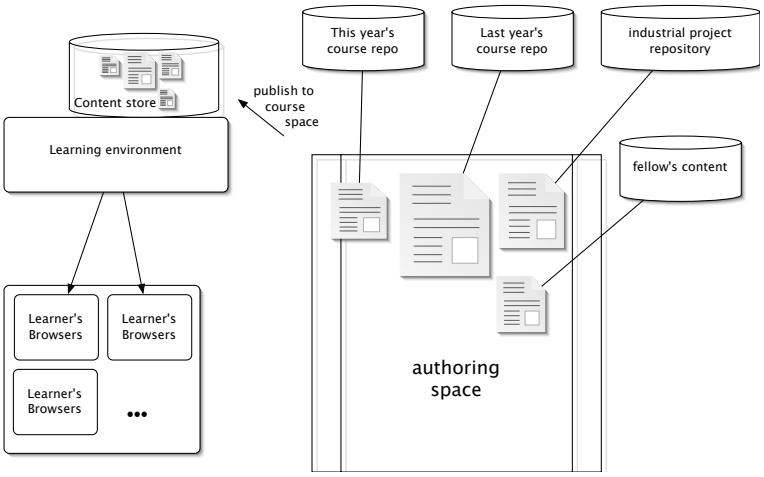

Fig. 1. Projects' organizations of the author creation process of the user-story.

Re-use sheds a new light on these spaces since one such set of spaces is needed for each content project; therefore transport methods between these spaces is needed, for example from the originating author's publication space to the composition, staging, and publication spaces of the recipient authors. This paper articulates re-use under the perspective of the many relationships between authors and these spaces which are mostly all visible on the web.

Central to the re-use actions is the granularity of the content-items manipulated. The finer it is, the more management actions may be required but the more freedom for re-organization it allows; the coarser it is, the easier the content items are to exchange, but the higher the risk to require the creation of a new version. ActiveMath has chosen a very fine-grained approach where individual mathematical content paragraphs such as a definition, a motivation, or an example, are the content items. Most other tools have the granularity of a page (typical of content-management-systems or wikis), or of a book (typical of a desktop document). Items are combined into larger pages or books using the item-inclusion paradigm, that is, books are *assembled* (or *aggregated*) in table-of-contents hierarchies. This approach is what we call **item-inclusion**.

Another important aspect of re-use is the nature of the material and its content-encoding. The nature of the material varies from simple pictures to complete web-sites, the nature of the content-encoding varies from final-delivery executable materials (such as a multimedia CD or a video exploration), to the full authoring sources. Each of these dimensions impact strongly the re-use possibilities. For example, lack of authoring sources (or available software to edit it), means the sole possibility of acceptance or rejection *as-is* for re-use.

Because web-delivery technologies are changing rapidly, it is common for the authoring sources to be much more abstract. One way to abstract is to write **content semantically**. For ActiveMath this has meant that all formulæ are

encoded using the OpenMath standard and that all content-items are given a mathematical and pedagogical role (e.g. exercise, definition...). We refer to [4] for more details about the knowledge representation of ActiveMath. This gives the advantage that mathematical notations are homogenous since they are separated from the content.

Re-using a piece of e-learning content within another setting is a fundamental action to evaluate many of its facets. If the origins are sufficiently traced, it is possible to stimulate evolutions of the content pieces following this re-use and thus enable long-term **quality management**. Several such workflows which combine quality assessment, content composition, and usage have been studied in the e-Quality project, see, e.g. [5].

Last but not least, the **consistency** of the resulting learning experience is desirable for learners so as not to raise the load on extraneous memory. This fundamental aspect of the re-use result has to be proofed by re-using authors.

3 Current Practice in Re-using

In this section we attempt to classify the current practices that can be called re-use in e-learning. We catalogue the weaknesses and strengths of each.

Classical Re-Use: Verbatim Inclusion The simplest form of re-use is achieved by importing the file of someone else, not changing it, and including it within his authoring space. This practice is common, for example, for the re-use of pictures from clip-arts libraries. Verbatim inclusion works much less well as soon as changes are needed.

Classical Re-Use: Copy-and-Paste Most of the time, re-use is based on the usage of the *copy-and-paste* paradigm. The evaluation of the WINDS authoring tools states: *Reusability is not measurable in our system as it is based on the copy and paste method that can be applied on various levels – learning unit, material, content block and index.* [6]

Re-use through copy-and-paste can beof great help for imitation, which is an important activity to learn to author.

Re-use based on copy-and-paste and subsequent modification is clearly insufficient for the quality development. This practice generally looses the the trace of the inclusion (from where to where in the re-authored work). It thus prevents transmission of further enhancements.

Classical Re-Use: linking The simplest form of re-use in web-publishing is realized by the usage of a hyperlink from the published text into another.

Linking is a method that is very easy for authors and keeps good tracks of the origin (hence can take advantage of enhanced versions). However, when followed by learners, it looses almost fully the users' context within the learning environment and the practing of sending learners to uncontrollled sites has attracted mistrust as in [7].

Classical Re-Use: Copy and Branch of Large Bodies The operation of *branching* is, typically, achieved when moving from one project to another by duplicating the document then arranging the various parts. The branching practice could be best of breed in principle but needs, in general, to be combined with further edits: for example, re-using a textbook chapter or the slides of a whole lesson will very likely lead to modifications. Analysis of differences so as to flow back and forth enhancements, as described in the user-story above, may become difficult with such a practice because of the grain size. Indeed the usage of a differencing tool is then required but rarely practical.

Classical Re-use: Channel-based as in News-Feeds One of the common practices of web-based re-use is the practice of newsfeed aggregation where one *aggregates* news from several sources and merge them in one news-page. The resulting stream of news includes both locally created news texts as well as remotely fetched news-texts. They are generally presented in a chronological manner and serve targets such as web-logs or groups' news pages. This paradigm is most commonly exercised by the abonnement to RSS-channels, which are simple URLs of RSS or Atom documents, two XML formats which describe a timed sequence of news items. Channel based re-use is very easy to activate: one needs only drop the URL into the configuration of the aggregator. Channel based re-use is not appropriate for e-learning contents and the same unpredictability as linked re-use is there, finally modifying an included news item is not a normal practice.

4 Project- and Inclusion-Based Re-Use

In the previous section, we surveyed widespread re-use practice and saw their strengths and limits. In this section, we propose a model to content sharing and re-use for e-learning that appears to gather advantages of each method described above. To our knowledge, this model is new although it seems to be closely implementable in Connexions (but see section 9.1).

We propose to organize re-use along *content projects*, which we shall call, in ActiveMath, *content collections*. More or less they are created at organization time, when a content goal is formulated and have a lifetime as long as the evolution of the intended learning experience (which may span several years). Content projects have a broader scope than single books. An example would be described in the user story in section 1 as the project of industrial examples.

To realize a re-use action of a content project, the learning environment using them needs to define the ability to *import* items of a content project, that is the actions to bring within the scope of includable content another content collection. Such an action does not mean the inclusion of the whole content within the current authored work. Further, we propose that learning environments allow imported content to be *included* within the realization of another content collection, that is its content-elements become part of a larger entity, for example content items become part of a *book* navigation artifact. Other actions with content projects include the browsing evaluation, download, appropriation, and publication, all of which are tasks of an authoring support tool.

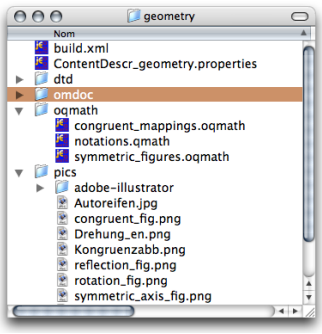

Fig. 2. A file-explorer view of a typical content collection:

ActiveMath's content collection implement this pattern: to the eyes of the ActiveMath learning environment, all content elements are contained in a collection, they are directories that contain a descriptor, server-content-files, as well as static web-resources. The screenshot in figure 2 depicts a file-explorer view of a typical collection ActiveMath: the content files in the `omdoc` directory, the authoring sources of them (the `oqmath` files), the current syntax-specification (`dtd`), a content-descriptor, and a directory of pictures (`pics` and their sources `adobe-illustrator`). This collection contains the enough information for a *build process* to be run which will compile the sources and upload them to the server as is described in [8].

Content collections should, at best, be focussed on a theme, they should have a long lifetime and carry a community of authors and consumers interested in it; this is strongly different than single documents shared in learning-objects'-repositories.With such a long lifetime, quality and evolution can be managed at the project level and different views can be exchanged about the usage of the content. As a result, content projects may come with their own versioning repository, their own download space, their own web-pages and their own community space.

Designating a content project as the **contents of a shared directory**, without having the need to move them around, thanks to the inclusion mechanism, allows the proper management of content-items while keeping a common file organization shared between contributors of the content project. Thus, the recipient of a re-use action is able to identify that the files he changes are coming from a project shared with some other people to which he may be able to offer his modifications. This location tracking is precisely what is missing in the copy-and-paste or the copy-and-branch practice which lead to an unmanaged proliferation of almost duplicates.

Advantages of the linking paradigm of re-use remain in a learning environment that allows reference-based inclusion. Indeed, the inclusion within a book's table-of-contents is a single line in ActiveMath's OMDoc. This inclusion is stronger than linking since it really inserts the content-item within the scope of the other items. Moreover the fine granularity as well as a semantic nature of contents help to make this re-organization consistent since they allow, each, paragraph level re-use or modification, (mixing reused and own items on a page) as well consistent math-notations even though the source is of a fully different origin.

5 A Simple First Approach: a Shared Authoring Server

It has been suggested several times that single global server, such as that of Wikipedia or Connexions would allow the re-use story of section 1) to be easily realized. The biggest issue of this approach is the impact of a content change. There is no reason that a given content author changes a Wikipedia page, or any other *central server* page, for the purposes of his intended teaching experience alone; Wikipedia pages are meant for the humanity and not for a particular classroom or book audience.

In the ActiveMath learning environment, for example, the impact of a change can be quite radical: the addition of an exercise anywhere suddenly makes it selectable by its tutorial component [9], or displayable by the search tool. We see, thus, the (current) impossibility to manage appropriate *scopes* to restrict impacts of changes.

Finally, a shared server suffers the same issues as the linking re-use paradigm with respect to loss of context of the learners which have to leave their learning platforms for another.

The approach of a global server, however, has a clear interest as a content commons for the instructors and authors who wish to discover and share content items. The shared platform allows easy cross-linking and easy searching, two important facets of the re-use of new materials. This shared platform is what made the success of the Connexions project [10] which, however, does not try to provide an adaptive experience to the learners but helps the author realize a document that they can deliver to the learners within a different serving environment. Most probably, domain-oriented author communities are the best host to such content-commons but the need to *export* from a content collection on a content commons to one's own learning environment remains.

6 How can an Author Re-Use Content?

Having described the spaces where the content projects and their sources can be stored, we now try to concretize the user story with successive approach steps an author may have towards the re-use of a content project. They are connection relations between an author and a content collection.

6.1 Remote Browsing

Remote browsing is the simple relation of the author's web-browser to be a client of a learning platform server running anywhere in the world. Such a connection relation is appropriate for discovery but requires the author to understand what is content in the web-experience he is having in order to know what he will be able to re-use. The author also cannot see how a content he would re-use would be integrated with other content items. Having identified the content items of interest, he is able to identify the content project he wishes to be re-using and can enter one of the connections relations below.

6.2 Web-Service Content Channel

Similarly to RSS channels, the idea of this connection method to a content storage is to use a simple remote protocol as single method of access to the content. Using this approach the author's learning-platform, for example an authoring-preview server, is configured to request anything concerning the content collection through this channel. In learning platforms, there seems to be few implementations of such a method.

ActiveMath implements this approach currently using the bare-bones XML-RPC protocol, a simple and efficient remote method invocation mechanism through *http*. It has been tested successfully, reaching 300 queries per second.

The web-service approach works fine for a small amount of queries and it is very easy to add a new collection, since it simply requires to add the URL of the web-service to the configuration.

6.3 Download of the Collection

A simple way to load a collection which avoids all network fragility is to simply download the collection in full, for example as an archive or through a *checkout* of a versioning server. Following the configuration of the learning platform to load this content collection, it needs to be *loaded*, that is, the *index* for this collection needs to be built which may be time consuming. This connection relation, being lengthy to set up, is not appropriate for a simple discovery of the content but provides the biggest safety. Also, it opens the door to appropriations:

6.4 Appropriation of Collections

It has been our experience that re-use is only acceptable to most authors if the right to modify is granted as well (the derivation right in [11]), along with the rights to redistribute (at least to one's learners) the modified content. The necessity to support the modification of content collections that are re-used is important. We call appropriation, the action of starting to modify a collection that is re-used since it is the action of *making own* this content collection. To our knowledge, appropriation only works with a downloaded copy of the content-collection and is thus an action only possible when one has really decided to try to re-use a content-collection. Once appropriated, the content collection would best be re-published, in order to be offered for further re-use:

7 Sharing of Content Project on the Web

In the previous section, we have described three methods above to access a content and its storage through the web. We now turn to describing the essential ingredients required for a publication on the web so that other authors can easily re-use a collection (even an appropriated collection).

As fundamental ingredient of a publication on the web is a single identifier which should not change. The first role of such a string is to be used as an

identifier to let the content storage identify its items, as well as let other content items reference it. The second role is used in discussion about collections and items. It is important that people may refer to a collection or a content item anywhere in electronic communication. The third role, if care is taken that this identifier is an HTTP URL, follows the *namespace document* practice of [12].

Here is information we expect to be found behind such a URL:

- a license that specifies the rights for re-use
- a list of names of copyright holders
- identifiers of the content projects this project relies on
- a list of all the resources involved in the collection
- list of the various storage connection configurations which allows the collection to be re-used
- optionally links to a community space of the content development, to courses or other events that make use this content collection, to usage reports, ...

All this information can be encoded within IMS Content Packaging manifests [13] (IMS-CP) which, by its XML nature, may provide this double role of machine processable file as well as browser-renderable document.

With such an information behind a content-collection manifest an author could be able to simply drag-and-drop the link with this URL into his authoring interface's configuration in order to let his authoring platform suggest connection methods and *be connected* to this content collection to use it.

8 A Web of Versions for Peer Productions

Creating a new version of a content collection published on the web can be done in two ways:

- A minor and compatible version change should probably be published at the same location with simple updates of the storage. Simple web modification tracking, similar to bookmark tracking, can be used to detect changed manifests, indicating a changed content.
- A major version change, for example between the course of one year to another, should probably best be made as a branch, allocating a new collection URI. Such a version change is the same as building and publishing an appropriated collection with a few modifications. Licenses that allow derivation allow this to be done by anyone in the world, who can choose to publish on the web an archive of the appropriated content collection.

The web of versions has content-package-manifest as nodes and dependencies or new versions between collections as links. In this web of version, a very important activity is the navigation between the various collections while shopping for the appropriate content to be re-used. For this, the ability to access a content-collection as easy as a link-click is very important. In this navigation, access may mean simply *browse about*, that is to read the small information about it, to try

it in preview servers (linked from this information), or to start using it in one's own platform.

Updating a content collection can be done in several ways: if it is a minor version change, an update or download is only needed. A major version change yields a change in URL: the new content collection needs to be connected to, the old version probably needs to be removed, and updating references to this collection means changing the collection-identifier.

To help the management of such identifiers, the OMDoc language used in ActiveMath has integrated the *development graph* [14]: this namespace management methodology requires that each document be endowed with *imports* elements which provide the full collection name of *theories* that are referenced. References within the document can thus use only the short *theory-name*. Major updates of relations, whereby one changes the URL of a content-collection, are done by only changing the imports. Because imports are recursive, this needs to be done once per content collection.

9 Conclusions

In this paper, we have presented approaches to organize content projects, to publish, to store, and to serve them, in order to facilitate re-use of content and long term quality management. Our approach is unique under several perspectives which we describe.

9.1 Originality of the Approach

We propose to rely on a very fine granularity of content items, which we suggest be as small as a motivation, a definition, or an exercise and be used as elementary units of inclusion. Opposite to this fine granularity we propose content projects which carry connection methods, ownerships, and authoring practice to be large and long lived.

This decoupling seems, in particular, to be lacking in all learning object repositories and many approaches that speak about learning objects where re-use is too often considered at the individual document level without a notion of a package derived from another.The large granularity of the projects may be key to the constitution of the communities' focus which, otherwise, may become scattered around too many resources. Both [5] and [2] indicate the community constitution as essential factor for re-use to happen.

Our approach considers the appropriation tasks deeper than [3] which proposes limited re-factoring methods activated by menu entries; instead, our approach insists that arbitrary modifications should be authorable in the appropriation process and let the differencing and versioning tools act, if need be, to identify or visualize the modifications. This task is best served, as of today, by source formats using source versioning tools.

The Connexions content common, seems to be closest to achieve our user-story. Although it provides combination and branching, it seems not to support

the *merging* or *return* of adaptations to help in quality management.[1] Another major difference to ActiveMath is the granularity of the item of content reorganization: Connexions' modules seem to be rather expected to be full pages making them harder to combine.

The ability for *anyone* to publish a new version is provided legally by licenses that allow derivation and technically simply by putting it on the web with a description of its source. This facility opens considerably the door to contributions for anyone with a web-space who can publish what he has considered appropriate for his context and usage.

The decentralization and distribution facet is slowly appearing in projects which re-use Wikipedia material: indeed the needs of school for a predictable, high-quality material has justified such an institution as SOS Children to fund the delivery of an offline set of wikipedia articles in a DVD form.[2]

9.2 Open Questions

Aside of completing implementation and large scale evaluation of our approach, open research questions remain:

It appears that, technically, the web-service channel approach could be relaxed to exploit simple web-of files using a process of compilation and map of XML content. This avenue could enable authors to use their normal web-space to publish their content-collections by a simple upload (such as FTP) and let other authors enjoy the same *lightweight installation* that the web-service channel offers, with local download of "only the files needed" for appropriated items to be edited.

The right to derive in public only: Our experience has shown that authors that invested body and soul in the creation of content find it hard to let anyone change their work freely and redistribute it even though the license stipulates well the requirement to quote the original author; the most common justification is that authors wish to *know what will happen to their content.*

A potential avenue to solve this is to request the *notification of usage and change*. Global web-availability of the displayed content with links to the collections' identifiers may be sufficient to answer this: reverse links queries, as supported by current web search engines, may reveal the usages; it would enable an author to see where a content project is used, to browse it, and, if need be, to request the removal of his authorship. A mix of legal and technical investigation is required in order to make the derivation right less feared about.

The Right Brick of E-learning Peer-Production? Yochai Benkler in [15] analyzes the ingredients that have allowed the peer productions spaces to exist and grow as large as we know them (e.g. open-source software or the Wikipedia initiatives). Among the major ingredients are the possibility of a very fine grain

[1] A private communication with authors of the Connexions project seems to indicate a growing feature-set that allows fine-grained combination (through the course composer) and management branches and differences.

[2] See `http://schools-wikipedia.org/`.

contribution and its affordance. We claim that the contribution of a delta posted on a static web storage satisfy these requirements and thus believe that this framework can allow decentralized peer-production.

References

1. Robson, R., Collier, G., Muramatsu, B.: Reusability framework. Technical report, Reusable Learning Project (2006) See http://www.reusablelearning.org/.
2. Busetti, E., Dettori, G., Forcheri, P., Ierardi, M.G.: Promoting teachers' collaborative re-use of educational materials. In: Innovative Approaches for Learning and Knowledge Sharing, Proceedings of ECTEL-06. (2006) See http://www.springerlink.com/content/n25p524n4514/.
3. Meyer, M., Hildebrandt, T., Rensing, C., Steinmetz, R.: Requirements and an architecture for a multimedia content re-purposing framework. In: EC-TEL. (2006) 500–505
4. Melis, E., Büdenbender, J., Goguadze, G., Libbrecht, P., Pollet, M., Ullrich, C.: Knowledge representation and management in activemath. Annals of Mathematics and Artificial Intelligence, Special Issue on Management of Mathematical Knowledge **38**(1-3) (2003) 47–64 Volume is accessible from http://monet.nag.co.uk/mkm/amai/index.html.
5. Montalvo, A.: Conceptual model for odl quality process and evaluation grid, criteria and indicators. Technical report, e-Quality Project (2005) See http://www.e-quality-eu.org/.
6. Kravcik, M., Specht, M., Oppermann, R.: Evaluation of winds authoring environment. In: Proceedings of Adaptive Hypermedia and Adaptive Web-Based Systems, AH 2004, Eindhoven, LNCS 3137. (2004)
7. Moravec, J.: US senator: Ban wikipedia from schools (2007) Available at http://www.educationfutures.com/2007/02/15/us-senator-ban-wikipedia-from-schools/.
8. Libbrecht, P., Gro, C.: Experience report writing LeActiveMath Calculus. In Borwein, J., Farmer, W., eds.: Proceedings of Mathematical Knowledge Management 2006. Number 4108 in LNAI, Springer Verlag (2006) 251–265 Available at http://www.activemath.org/pubs/bi.php/Libbrecht-Gross-Experience-Report-Authoring-LeAM-MKM-2006.
9. Ullrich, C.: Course Generation as a Hierarchical Task Network Planning Problem. PhD thesis, Saarbrücken (2007)
10. Henry, G., Baraniuk, R.G.: Peer to peer collaboration using connexions. In: American Society for Engineering Education Annual Conference & Exposition Salt Lake City, Utah. (2004) Available from http://cnx.org/aboutus/publications.
11. Creative Commons: Choose a license (Accessed in April 2008) See http://creativecommons.org/license/.
12. Jacobs, I., Walsh, N.: Architecture of the world wide web. Technical report, World Wide Web Consortium (2004)
13. Smythe, C., Jackl, A.: IMS Content Packaging Information Model v1.1.4 (2004) See http://www.imsglobal.org/content/packaging/.
14. Mossakowski, T., Autexier, S., Hutter, D.: Development graphs – proof management for structured specifications. Journal of Logic and Algebraic Programming **67**(1-2) (2006) 114–145
15. Benkler, Y.: The Wealth of Networks. Yale University Press (2006) See http://www.benkler.org/wealth_of_networks/.